

学校编码: 10384

分类号_____密级_____

学号: 19020111152543

UDC_____

厦 门 大 学

硕 士 学 位 论 文

Theil系数和主成分聚类分析在区域经济差距评估中的应用

Application of Theil Coefficient and Principal Component Cluster Analysis in Assessment of Regional Inequality

张 志 标

指导教师姓名: 谭忠教授

专 业 名 称: 应用数学

论文提交日期: 2014 年 月

论文答辩时间: 2014 年 月

学位授予日期: 2014 年 月

答辩委员会主席: _____

评 阅 人: _____

2014 年 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ） 1.经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

（ ） 2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

摘要

由于中国三十几年的改革开放，总的来说，中国的经济实力和人民的生活水平有了很大的提高。当然，与发达国家相比还是有很大的差距的，我们还有很长的路要走。但是人们慢慢地发现在区域间的经济发展极不均衡，使得人们的收入水平有很大的差距，高低不一。然而，在1990年之后，学者们才真正对这种问题进行了分析。近十年来，这个问题普遍受到各阶层的关注。

大多数学者们进行经济差距探讨主要是通过不同地区的划分，像省为单元、三个地带等。研究的方法也不同，有的人选择多个指标形成指标体系对区域经济进行综合评价，有的人使用国际上惯用的指标，比如*Theil* 指数、*Gini* 指数。然而，迄今为止，人们尚未形成应该如何度量区域经济差距的相同意见。因此，他们的研究结果也是不相同的。

锡尔系数对富裕地区的收入变化反映较快，即锡尔系数的变化情况与富裕地区收入变化是成正比的。当选择多个指标探究区域经济差距时，考虑到选择的指标较多，这样有些指标可能包含着有一部分相同的信息，从而它们间的相关性也大。要是能通过某种指标间的重组，形成较少的指标又没有损失太多的信息，这时主成分法成了最佳选择。当我们要对目标的性质进行分析比较时，仅靠主成分是不够好的，况且它多少还是损失了一点信息，这时聚类分析起到较好的作用，它可以对主成分分析和*Theil* 系数进行补充。

本文中，将融合聚类分析、*Theil*系数和主成分分析的优点。先用*Theil*系数对我国2001 – 2010年期间的三个地带间的经济差异变化过程进行实证，然后在主成分的基础上进行聚类以对2010年各行政区域的经济差异进行实证研究。经过分析得出的两个子结果后，本文得到的结论是地带间的差异在整体差距中起决定性作用，而沿海与内陆的经济差距又是地带间的主要差距。

关键词：*Theil*系数；主成分聚类；区域收入差距.

Abstract

Generally speaking, China's economic strength and people's living standards have been greatly improved since the implementation of reform and opening up more than 30 years ago. However, compared with developed countries, we still have a long way to go. As China's economy continues to develop, people gradually find that the rich become richer, and the poor become poorer. In other words, the problem of disparity is very serious in China, which didn't draw much attention until 1990s. However, the recent decade has seen an increasing attention from all walks of life paid to this problem.

Most scholars researched into the issue of disparity by dividing the country into different regions, such as administrative regions and the three-belt division. So far, researchers haven't reached an agreement on the research methodology of regional economic disparity. Some of them use indicator system, and some others select internationally used index such as Theil coefficient. Naturally, the research results are different due to different methodologies.

The Theil coefficient is sensitive to the income change in developed regions. That is to say, the more the income in the developed regions increases, the bigger Theil coefficient becomes. When we evaluate regional economic disparity, we may use many different indicators. These indicators often contain similar information and have great relevance. To avoid this situation, we make full use of principal component analysis (PAC) producing new indicators consisted of raw indicators via linear combination. Although PAC may cause a little loss of information initial indicators include, it is convenient for us generate less new indicators including most of raw information. However, when we compare the features of the subjects, PAC is far from enough and it causes a little loss of information. At this time, cluster analysis should be introduced to complement the function of PAC and Theil coefficient.

This paper makes full use of the advantages of the three methods - Theil coefficient, principal component analysis and cluster analysis. On the one hand, the paper analyses the changes of income disparity in the three belts of China from 2001 to 2010 by using Theil coefficient. On the other hand, it explores the income disparity of all the administrative regions in China by means of principal component and cluster analysis. After analyzing the study results results obtained from the above two methods, the paper concludes that the income disparity among the three belts of China plays a decisive role in the income disparity among the country's administrative regions; what's

more,the income disparity is most obvious between the eastern belt and the central-western belt of China.

Key Words: Theil coefficient; Principal component cluster analysis; Regional income inequality.

厦门大学博硕士论文摘要库

目 录

摘要	I
Abstract	III
第一章 绪 论	1
1.1 选题背景与意义	1
1.2 相关文献综述	2
1.2.1 国外区域差距文献评述	2
1.2.2 国内区域差距文献评述	4
1.3 本文的主要工作	5
第二章 主成分聚类分析	6
2.1 降维的基本思想	6
2.2 主成分的基本概念和结论	8
2.2.1 主成分的基本性质	10
2.2.2 关于主成分的计算过程	12
2.3 聚类分析的基本思想	13
2.3.1 距离	14
2.3.2 相似系数	15
2.4 系统聚类分析	16
2.5 动态聚类分析	17

2.6 聚类收敛性分析	19
第三章 区域差距评估指标框架	21
3.1 选择评估指标的原则	21
3.2 国外常用指标	22
3.3 国内常用指标	24
第四章 我国区域差距的实证研究	27
4.1 Theil系数的应用.....	27
4.2 主成分聚类分析的应用.....	29
4.3 实证结果分析	30
结论	36
附录 A	38
参考文献	41
致谢	43

CONTENTS

Chinese Abstract	I
Abstract	III
Chapter 1 Introduction.....	1
1.1 Backgrounds and significance of the study	1
1.2 Literature review	2
1.2.1 Literature review of regional inequality at abroad.....	2
1.2.2 Literature review of regional inequality at home	4
1.3 Main works of the thesis.....	5
Chapter 2 PCA and cluster analysis	6
2.1 The main idea of dimensionality reduction	6
2.2 Basic concepts and Theorem on principal component analysis ...	8
2.2.1 The properties of principal component analysis	10
2.2.2 Calculation	12
2.3 The basic idea on cluster analysis	13
2.3.1 The distance	14
2.3.2 The Similarity coefficient	15
2.4 Hierarchical cluster analysis	16
2.5 Dynamic cluster analysis	17
2.6 Cluster Convergence	19
Chapter 3 Indicators to estimate regional inequality	21
3.1 The principal to select indicators	21
3.2 Indicators used at abroad	22
3.3 Indicators used at home	24
Chapter 4 Empirical Study of our regional economy.....	27
4.1 Application of Theil coefficient	27

4.2	Application of principal component cluster analysis	29
4.3	Analysis of the empirical result.....	30
Conclusion		36
Appendix A		38
References.....		41
Acknowledgements		43

第一章 绪论

1.1 选题背景与意义

经济差异一直以来都是一个热点话题，因为它的出现与许多重要的问题相关的。不仅仅是在经济学界，而且还涉及社会、政治和其他诸多方面。在这个意义上，*Sen*(1973)提出了:在多大程度上的经济差异会与使社会不稳定的因素联系在一起的。实际上，这包含着社会的动乱。

当然，它也意味着其他一些有趣的经济概念，比如说收敛性、发散性、贫困、福利等。因此，近二十年来，一些学者们把大量的精力放在经济差距方面。但是最早研究经济差异的应该是*Atkinson*(1970)和*Sen*(1973)，他们的研究集中在诸如差异的量化、经济差距形成原理和差距间的比较等方面。

当我们想起区域经济差异时，人们通常的一个想法是地区间的平均收入差距。也就是说，我们用每个区域的平均水平去替代该区域中每个居民的实际收入。虽然这个概念能有效表现出该区域的整体经济发展实力和进程，但是它却不能很好地探讨人们的收入不同是否是因为他们属处的地理位置不同或者是否是因为在相同的区域有着较大差距的收入。在这种情况下，区域差距的分解就成为对区域经济差异进行积极分析和社会政策设计和评估的一个有用的工具。

中国从1978年开始实行改革开放，产生了让世人震惊的成效。与计划经济时代相比，不仅国民经济实力显著增强，到2013年我国的国内生产总值达到了564916.25亿元，排在世界第二，仅次于美国(1012372.5亿元)。这样的发展速度，放眼世界可谓前无古人，让很多国家大为惊讶，特别是让拥有相同制度国家无法望其项背。

但是，我们不能安于现状，如果我们冷静细思，我国国土面积960万平方公里、人口已超过13亿，这些数字在全世界都是位列前茅的。我国2013年的人均生产总值是5414美元，位列世界89，排在世界第一的是卢森堡113533美元，第九的加拿大50436，第13的新加坡49271美元，第14的美国48387美元，第十八的日本45920美元，第十九的法国44008美元，第二十的德国43742美元，第二十二的英国38592美元，第三十五的韩国22778美元。

从上列举的数据知，他国的人均GDP要远远高出我们很多。从另外一个方面来说，虽然我国内总产值很高，位列第二，但是人均总产值很低，排到了89

位，这说明了我国的区域发展是极不协调的，是有差距的。当然，我们不能因为存在的差距，就否认中国数十年来取得的巨大成就，就否认改革开放的政策，因为人民的日子过得比以前好了，这是不容置疑的。

翻阅发达国家的历史，我们会发现它们在发展经济的过程中，不同的区域，人民的收入水平是不相同的，是有差异的。即要发展经济，区域间的差异就客观存在，而且是由发展而产生的，不可能每个区域都同步发展，这已经被历史所证明。

然而我们也不是在差异面前无能为力的。虽然我们不能消除因为要发展而存在的差异，但是我们却可以使这种差距缩小。30年代美国的罗斯福新政就是一个很好的例子。像党所提出的：我们的发展是为了人民的发展。

虽然人们比以前富裕了，但是收入悬殊、区际间差距大、各种资源分配不合理、社会的各种不良因素增多。人民对公平、公正有着更高的期望。

要改变所出现的问题，我们就必须客观地正视问题。差异的存在是毋庸置疑的，但是差距到底有多大，现在具有的差异是不是已经超过了合理的范围？我国存在的经济差异是什么样的差异？是三地带之间的差异？还是城镇与乡村间的差距？亦或是省级层次间的差异？当我们找到了这些矛盾的根源，方可设计出合理、科学的政策去缩小这些差异，才能更好的发展国民经济。

1.2 相关文献综述

1.2.1 国外区域差距文献评述

区域差距的研究在西方国家已有了较长的一段历史。既有关于理论方面的，也有对不同区域的定量的实证探究。简要地概述如下：

起初，当经济差距逐渐呈现在人们的眼前，特别是对于研究者而言，他们会不禁问道收入上的差距在未来是发散还是收敛？于是，就产生了能否均衡发展的激烈探讨。

认为区域发展会趋于均衡的代表人物有J.L. 斯坦等人，他们认为像诸如人力、资本、技术等生产要素会在不同的地方自由移动，也就是说生产要素在各个地方的分配最后也是会趋于平衡的，这样自然就会逐步缩小不同地区的贫富差距，进而实现区域的同步发展。

但是瑞典的著名经济学家缪尔达尔等人却持有截然相反的观点，他们认为区

Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”. Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库